Noname manuscript No.
(will be inserted by the editor)

Non-intrusive Quantification of Performance and its Relationship to Mood

Davide Carneiro · André Pimenta · José Neves · Paulo Novais

Received: date / Accepted: date

Abstract The number of jobs that takes place entirely or partially in a computer is nowadays very significant. These workplaces, as many others, often offer the key ingredients for the emergence of stress and the performance drop of its long-term effects: long hours sitting, sustained cognitive effort, pressure from competitiveness, among others. This has a toll on productivity and work quality, with significant costs for both organizations and workers. Moreover, a tired workforce is generally more susceptible to negative feelings and mood, which results in a negative environment. This paper contributes to the current need for the development of non-intrusive methods for monitoring and managing worker performance in real time. We propose a framework that assesses worker performance and a case study in which this approach was validated. We also show the relationship between performance and mood.

Keywords Worker Performance · Mood · Statistical Analysis · Distributed Intelligence

1 Introduction

We currently live in a society that moves at an unprecedented pace. People need to find a balance between work, family and leisure, and doing so is increasingly harder [7]. In an attempt to achieve it, people stretch their limits and their days,

Davide Carneiro
CIICESI, ESTG, Polytechnic Institute of Porto
Porto, Portugal
E-mail: dcarneiro@estg.ipp.pt

André Pimenta, José Neves and Paulo Novais
Algoritmi Centre/Department of Informatics
University of Minho
Braga, Portugal
Tel.: +351 253 604 437
Fax: +351 253 604 471
E-mail: apimenta,jneves,pjon@di.uminho.pt

usually at the expense of rest and relaxation time. This, together with inadequate sleep patterns are leading causes of fatigue.

Mental fatigue may not be felt immediately or may not even be felt at all, but it has effects at many different levels, including health, performance, decision-making and even safety. These effects are often visible in students preparing for exams, office or industrial workers, health care professionals, drivers, pilots or military personnel. Fatigue may even put people working in safety-sensitive jobs at risk, and any mistake on their part can lead to loss of lives [17, 6].

Mental fatigue can be seen as a state that involves a number of effects on a set of cognitive, emotional and motivational skills and usually results in overall discomfort, as well as the emergence of performance limitations [11]. Some of these limitations imply that a tired person is often less willing to engage in tasks of effort, or perform the task of a conditioned form, well below their normal capacity [8]. Thus, mental fatigue may be characterized by a perception of a lack of mental energy. Persons who are affected by mental fatigue may feel like they have less energy than usual and are unusually tired and lethargic. Excessive activity and stimulation of the brain can cause a person to feel mentally exhausted, and the feeling is similar to what the body feels when a person is physically fatigued [18].

Fatigue effects may occur at any moment and they may persist from only a few hours to several consecutive days. Depending on its duration and intensity, fatigue may make the carrying out of daily tasks increasingly difficult or even impossible. In severe or prolonged cases, it can cause illnesses such as depression or chronic fatigue syndrome [17]. The importance of detecting and managing fatigue, especially in the workplace, assumes thus an unquestionable importance nowadays.

This paper presents a framework for monitoring worker performance in real-time and in a non-intrusive way. For this purpose, the framework acquires and analyzes worker performance indicators. Specifically, it considers the performance of the interaction with the computer: how the user interacts with the mouse and the keyboard. We describe a case-study in which this framework was used to assess the performance of two groups of students: one taking classes in the morning and the other in the afternoon. We show that performance is significantly lower in the afternoon. We assessed the participant's fatigue in two different ways: through a questionnaire and using our framework. Both approaches achieve very similar results, which point out increased fatigue in the afternoon. Moreover, the afternoon group also evidences more negative states of mind.

The proposed framework may thus constitute an important step not only for monitoring and managing performance in the workplace but also to improve other inter-related aspects such as mood, well-being and workplace quality [3].

1.1 Profiling and Others Factors

There are many factors, both internal and external, that modulate the onset and presence of fatigue. These include sleep deprivation, naps, noise, heat, mood, motivation, time of day, and workload, and of course the individual profile [2].

The user's profile provides valuable information with respect to the potential level of fatigue. It can be seen as a predicted base level of fatigue in the sense that it establishes a baseline, according to the lifestyle of the individual [1]. These aspects have been thoroughly studied, mostly by psychologists, and encompass:

- **Age** - Defines the mental age of the individual. It is important to understand the expected cognitive abilities of the individual, which may have a tendency to degrade with old age.
- **Gender** - The mental states are different between men and women.
- **Professional occupation** - Many occupations are intrinsically more tiresome or exhausting than others.
- **Consumption of alcohol and drugs** - The use of certain substances for short or prolonged periods may cause dependencies and other effects that lead to a state of mental fatigue.

Mental fatigue is also affected by a number of other external factors. They may or may not be directly related to the individual's behavior. They include:

- **The mood** of the individual may influence decisively his or her mental state, with a particular effect on his or her motivation to work. Although tired, the individual may overcome (even if only temporarily) the effects of fatigue with a positive mood and motivation.
- **Stress** may be defined as the demands placed upon the individual's mind or body by external stimuli, requiring the individual to acclimatize to the dynamic requirements of the environment. However, these processes of acclimatization require an additional effort from the brain which, when prolonged over long or intense periods, will result in mental fatigue.
- **Mental Workload** as a result of the relationship between the amount of mental processing capacity and the amount required by the task.
- **Sleepiness** is often mistaken for mental fatigue or generalized as such. A difference exists and must be pointed out. However, the mistake is understandable since sleepiness is a symptom that is strongly connected to mental fatigue: it is one of the methods our brain uses to tell us that it is running out of resources. Sleepiness often results in a general loss of the individual vitality.

These factors can be assessed using validated tools such as the USAFSAM fatigue scale [15] for mental fatigue states, or as the NASA TLX [5] in the case of mental workload.

However, these instruments, based on the individual's subjective interpretation of the symptoms, do not fully take into account inter-individual differences. There are instruments that help to account for individual differences, such as the Profile of Mood States (POMS) [4] and the State-Trait Anxiety Inventory (STAI)[16]. However, their use in complex systems can prove to be complicated and confusing, due to the same problems that can be observed in subjective measures of fatigue detection [10, 4].

The POMS can also be used to estimate fatigue with the Vigor-Activity and Fatigue-Inertia factors. The POMS measures five aspects of affect or mood [9]. It consists of 65 adjectives describing feeling and mood to which the subject responds according to a five-point scale ranging from "Not at all" to "Extremely". Results are reported as six mood factors, namely:

- **Tension-Anxiety:** Heightened musculoskeletal tension including reports of somatic tension and observable psychomotor manifestation;
- **Depression-Defection:** Depression accompanied by a sense of personal inadequacy;

- **Anger-Hostility:** Anger and antipathy toward others;
- **Vigor-Activity:** Vigorousness, ebullience, and high energy;
- **Fatigue-Inertia:** Weariness, inertia and low energy level; and
- **Confusion-Bewilderment:** Bewilderment, muddleheadedness; appears to be an organized-disorganized dimension of emotion.

Because of its length, the POMS only results practical for occasional uses such as establishing a baseline and estimating the effects of excessive sleep deprivation or restriction.

2 Framework

In this paper we propose a non-invasive, non-intrusive, real-time approach to assess worker performance through the analysis of keyboard and mouse interaction patterns. The analysis takes place directly from the usage of an individual's computer as within the context of the so-called desk-jobs. We build on the fact that computers are nowadays used as major work tools in many workplaces to devise a non-invasive method based on the observation of the worker's interaction with the computer.

The main aim is to develop leisure and work context-aware environments that may improve quality of life, mental health and individual performance, as well as productivity in organizations. To achieve this purpose we follow the guiding lines of Ambient Intelligence (AmI) [14], in which the technological aspects are hidden in the environment and the user is placed in the middle of the paradigm. There is also a focus on non-intrusiveness, with acquisition of information taking place without the need for explicit or conscious user interactions.

Another major objective of the system is to support the decision-making processes of team managers or group coordinators. In this perspective, each element of a group/organization is seen as part of a whole which contributes to the general state of the group. Thus, the estimated performance level of the group is the average result of the fatigue level of its elements.

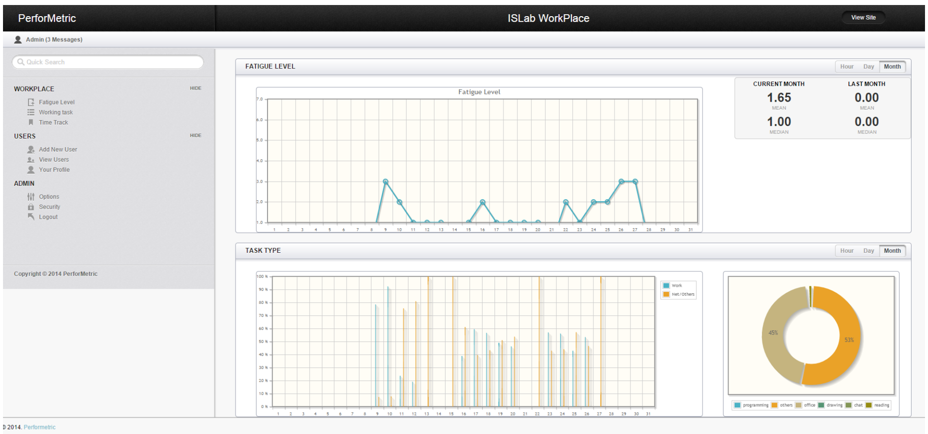

Fig. 1 Interface of the Performance Monitoring Web Service.

Specifically, the framework acquires and compiles data regarding the following features of interaction:

- Absolute Sum of Degrees - Quantifies how much the mouse moves or not in straight lines. A higher value in this feature means that the movement of the mouse has many curves, which is a sign of decreased efficiency, i.e., the movement is more efficient if it moves in a more straight line between each two clicks (units: degrees);

- Average Distance of the Pointer to the Line - Quantifies the average distance of the mouse pointer to the straight line defined by each two consecutive clicks. A higher value denotes that the user is moving the mouse farther away from this line, which is a sign of decreased efficiency (units: pixels);

- Average Excess of Distance - The distance of the straight line defined by each two consecutive clicks is the shortest possible distance between those two clicks. This feature measures how much the mouse pointer traveled in excess (units: pixels);

- Click Duration - It measures the length of each click, not considering long clicks associated to mouse dragging nor double clicks. The longer the click, the smaller the efficiency (units: milliseconds);

- Distance Between Clicks - This feature quantifies the amount of distance traveled by the mouse between each two consecutive clicks. A larger distance between each two consecutive clicks is a sign of decreased performance (units: pixels);

- Distance of the Pointer to the Line - This feature is similar to the previously mentioned Average Distance of the Pointer to the Line. However, it does not compute the average value but rather the sum of the distances of each position of the mouse to the straight line, between each two consecutive clicks (units: pixels);

- Excess of Distance - It is similar to the previously mentioned Average Excess of Distance, but the excess of distance is not divided by the distance between the two consecutive clicks (units: pixels);

- Mouse Velocity - It quantifies the velocity of the mouse by dividing the distance traveled (in pixels) between each two consecutive clicks by the time it took (in seconds) (units: pixels/milliseconds);

- Time Between Clicks - This feature quantifies the time, in seconds, that passes between each two consecutive clicks. Higher values are sign of a slower working rhythm (units: milliseconds);

These features describe, in an extensive manner, the performance of the user's interaction with the mouse. Additional features are also compiled by the framework concerning the use of the keyboard, window switching patterns, among others. These are, however, outside the scope of this work.

Besides the data acquisition module, the framework is composed of the following key components:

- Dashboard - An online reporting tool to aggregate and compile information about the team's performance, aimed at supporting decision-making and team management;
- Database - To maintain and manage the large amounts of information that are acquired and compiled about the workers;
- Machine Learning - A group of tools and algorithms to train models that characterize the users' interaction with the computer and the behavioral changes caused by different levels of performance.

3 Case Study

In order to test and validate the proposed system and approach, a case study was prepared using detection system models which were previously trained [13, 12]. To reach that goal we used two groups of volunteers who used the tool while they were providing feedback of their state of fatigue through USAFSAM Fatigue scale, as well as their mood through the POMS questionnaire.

Feedback values recorded during the monitoring period were used to validate the system and used to analyze the variation of moods.

The participants in the case study, fourteen in total (10 men, 4 women) were students from the University of Minho in the field of physics. Their age ranged between 18 and 25.

3.1 Methodology

The methodology followed to implement the study was devised to be as minimally intrusive as the approach it aims to support. Participants were provided with an application for logging the previously mentioned events of the mouse and keyboard. This application, which maintained the confidentiality of the keys used, needed only to be installed in the participant's computer and would run in the background, starting automatically with the Operating System. The only explicit interaction needed from the part of the user was the input of very basic information on the first run, including some personal identifying and profiling information.

As mentioned, two different groups of users were selected (7 participants in each group) who underwent the experience in different periods of the week. The first group was monitored on a Monday morning (starting at 9am), while the second did so on a Thursday in the afternoon (starting at 14pm). Each session had a duration of 3 hours. During the session, in addition to the collection of interaction features, user feedback regarding their state was collected hourly.

3.2 Data Analysis

A statistical analysis of the interaction data collected was carried out before analyzing the results in terms of mental fatigue classification. The aims of this analysis

Table 1 Mean and Median values of the distribution of the data for each feature and each group.

Feature	Mean		Median		p-value
	Group 1	Group 2	Group 1	Group 2	
Absolute Sum of Degrees	4786.89	7603.2	2716.49	3745.91	$5.15 * 10^{-27}$
Avg Distance Pointer to Line	37.08	49.63	15.89	23.01	$5.97 * 10^{-18}$
Avg Excess of Distance	4.07	8.02	1.46	1.78	$4.06 * 10^{-18}$
Click Duration	216.32	208.52	94.	94.	0.009
Distance Between Clicks	187.86	275.34	96.71	135.88	$1.28 * 10^{-17}$
Distance Pointer to Line	8390.94	16663.6	1352.74	2412.89	$2.00 * 10^{-20}$
Excess of Distance	340.73	612.51	83.04	167.44	$8.58 * 10^{-27}$
Mouse Velocity	0.42	0.41	0.18	0.18	0.003
Time Between Clicks	16579.7	17876.3	2282.	2547.	0.018

are twofold: (1) to demonstrate that there are, indeed, significant changes in the participants' interaction with the computer in different moments of the day; and (2) to demonstrate that the proposed framework is able to detect these changes based on worker performance.

Given the characteristic of the task attributed to the participants, that required mostly the use of the mouse, in this section we focus on the analysis of the features extracted from the mouse.

In this analysis, it is worth noting that both groups were carrying our similar tasks and that participants were randomly assigned to each group. The effects on performance that are shown here are thus potentially due to the effects of the circadian rhythm on the human body. Indeed, the performance of group 2 is significantly lower. This group carried out the assigned tasks in the afternoon.

Table 1 details, for each group and for each feature, the mean and median values as well as the p-value of the Mann-Whitney test, used to quantify the statistical differences of the distributions of the data for both groups. The mean values are higher in group 2 in seven of the features, while the median values are higher in seven of the features as well and equal in the other two. Moreover, all the distributions of the data are significantly different.

The data collected by the framework also allows more detailed and interesting analysis. For example, it is possible to analyze the evolution, over time, of any of the features for individual participants or for the group. Figure 2 shows the evolution of mouse velocity for two participants, one of each group. It is interesting to notice that at the beginning of the task both move their mouse with similar velocities (around 0.4 px/ms). However, as the task progresses, the evolution of mouse velocity is different for both groups: in group 1 it tends to increase (a sign of increased performance) while in group 2 it tends to decrease.

A similar trend can be observed in other features as well. Figure 3 shows the evolution of the excess of distance during the experiment, in the two groups. Once again, performance tends to increase during the morning and to decrease in the afternoon.

Similar trends can be observed in the remaining features. Essentially, the data shows two different and equally interesting aspects: (1) performance in the morning is significantly higher than in the afternoon and (2) these differences are statistically significant.

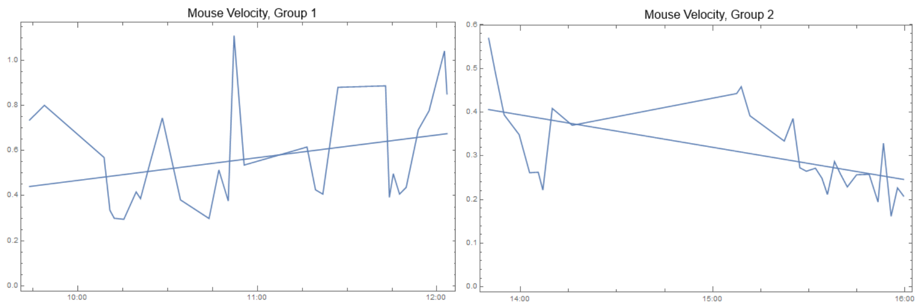

Fig. 2 Evolution of the velocity of the mouse during the experiment in the two groups.

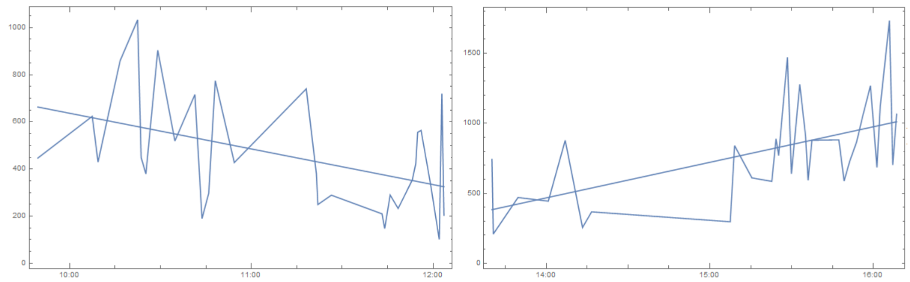

Fig. 3 Evolution of the excess of distance traveled by the mouse during the experiment in the two groups.

3.3 Results

During the two sessions the performance monitoring system was used in real time. It not only allows recording the interaction patterns with the mouse and keyboard, but also to calculate the estimated individual level of mental fatigue for each user, and the estimated average level of each group. These values were compared with the feedback given by the participants.

Table 2 summarizes the results, showing that the first group evidences better performance than the second one. This can be confirmed through data collected during the session. It is also possible to check a RMSE (Root-mean-square Error) of 0.2 for group 1 and 0.4 for group 2. Taking into account the scale used, this is an acceptable error.

Table 2 Average level of fatigue of the different groups in USAFSAM fatigue scale of the estimated values and feedback values

Group	Estimated Fatigue level (SD)	Fatigue from feedback (SD)	RMSE
1	1.6 (0.7)	1.4 (0.8)	0.2
2	3.3 (0.8)	3.2 (1.2)	0.4

In addition to the subjective level of fatigue through the USAFSAM fatigue scale it was used the POMS factors in order to observe the influence of fatigue on the moods of the different groups, and therefore the emergence of fatigue and loss of vigor.

Through Table 3 and Figure 4 it is visible that the average values are different between the two groups. In addition, the T-test was used to validate the differences between the two groups, in order to determine if the two sets of data are significantly different from each other, a fact that was confirmed.

We can also observe that fatigue is higher in the second group and, on the contrary, vigor is lower in the second group. This is in accordance to the fatigue levels of each group.

Table 3 Average vales of Profile of mood states for group 1 and group 2, as well as the resulting p-value from the T-test.

POMS factor	Group 1	Group 2	p-value
Tension-Anxiety	4.6	11.8	0.03
Depression-Dejection	4.4	9.1	0.02
Anger-Hostility	6.5	7.9	0.06
Vigor-Activity	19.4	7.6	0.04
Fatigue-Inertia	5.1	13.8	0.02
Confusion-Bewilderment	4.0	7.3	0.03

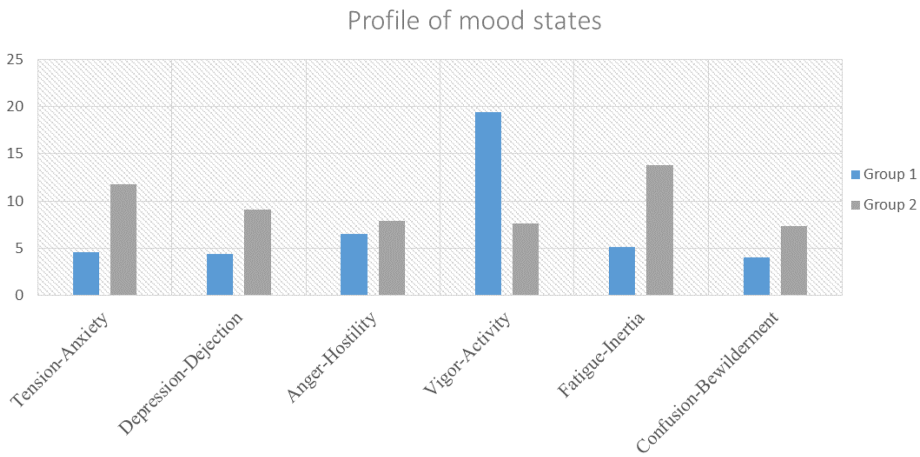

Fig. 4 Profile of mood states from Groups 1 and 2: group 2 has more negative mood states.

4 Conclusion

This paper described a framework for monitoring the performance of individuals or groups while they interact with the computer, in real time. As addressed throughout the paper, information about individual performance may be extremely useful in workplaces. Namely, it may allow a far better management of each worker, taking into account each one's working rhythms, context and state. This is especially important in scenarios of critical tasks, which require individuals performing at their best.

The results of the study carried out demonstrate that it is possible to evaluate performance in a group of people using their interaction with the computer as input. The prototype used in the experiment uses a model trained in a previous study, with data of a different group of users. This also shows that the causes of performance changes on different users can be generalized, i.e., we all react similarly. The study carried out also showed that in assessing worker performance and overall work quality there are other important aspects such as worker mood.

In the context of academic environments, such as in the case study, this may allow to define better schedules for classes or exams, for example, with the aim of improving student performance. In the workplace, similar strategies can be followed to improve worker performance. Moreover, and given that there are always inter-individual differences, personalized models can also be trained that shape each one's reaction. This will, undoubtedly, maximize the accuracy of fatigue classification and improve the success of performance management initiatives.

Acknowledgments

This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT – Fundação para a Ciência e Tecnologia within the Project Scope: UID/CEC/00319/2013. The work of Davide Carneiro is supported by a Post-Doctoral Grant by FCT (SFRH/BPD/109070/2015).

Compliance with Ethical Standards

Funding: This study was funded by COMPETE (POCI-01-0145-FEDER-007043) and by Fundação para a Ciência e Tecnologia (UID/CEC/00319/2013). Davide Carneiro has received research grants from Fundação para a Ciência e Tecnologia (SFRH/BPD/ 109070/2015).
Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.
Informed consent: Informed consent was obtained from all individual participants included in the study.

References

1. Akerstedt T, Knutsson A, Westerholm P, Theorell T, Alfredsson L, Kecklund G (2004) Mental fatigue, work and sleep. Journal of psychosomatic research 57(5):427–33, DOI 10.1016/j.jpsychores.2003.12.001
2. Balkin TJ, Wesensten NJ (2011) Differentiation of sleepiness and mental fatigue effects.
3. Carneiro D, Novais P (2014) New applications of ambient intelligence. In: Ambient Intelligence-Software and Applications, Springer, pp 225–232

4. Curran S (1995) Short Form of the Profile of Mood States (POMS-SF): Psychometric information. Psychological Assessment 7(1):80–83, DOI 10.1037//1040-3590.7.1.80
5. Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advances in psychology 52:139–183
6. Jaber MY, Neumann WP (2010) Modelling worker fatigue and recovery in dual-resource constrained systems. Computers & Industrial Engineering 59(1):75–84
7. Kobayashi H, Demura S (2006) Relationships between Chronic Fatigue , Subjective Symptoms of Fatigue , Life Stressors and Lifestyle in Japanese High School Students. School Health 2:5
8. Lorist MM, Klein M, Nieuwenhuis S, De Jong R, Mulder G, Meijman TF (2000) Mental fatigue and task control: planning and preparation. Psychophysiology 37(5):614–25
9. McNair DM, Lorr M, Droppleman LF (1971) Manual for the Profile of Mood States. Educational and Industrial Testing Service, San Diego CA
10. Newcombe PA, Boyle GJ (1995) High school students' sports personalities: Variations across participation level, gender, type of sport, and success. International Journal of Sport Psychology 26:277
11. Perelli LP (1980) Fatigue Stressors in Simulated Long-Duration Flight. Effects on Performance, Information Processing, Subjective Fatigue, and Physiological Cost. Tech. rep., DTIC Document
12. Pimenta A, Carneiro D, Novais P, Neves J (2013) Monitoring mental fatigue through the analysis of keyboard and mouse interaction patterns. In: Hybrid Artificial Intelligent Systems, Springer, pp 222–231
13. Pimenta A, Carneiro D, Novais P, Neves J (2015) Detection of distraction and fatigue in groups through the analysis of interaction patterns with computers. In: Intelligent Distributed Computing VIII, Springer, pp 29–39
14. Ramos C, Augusto JC, Shapiro D (2008) Ambient intelligence—the next step for artificial intelligence. Intelligent Systems, IEEE 23(2):15–18
15. Samn SW, Perelli LP (1982) Estimating aircrew fatigue: a technique with application to airlift operations. Tech. rep., DTIC Document
16. Spielberger CD (2010) State-Trait Anxiety Inventory, John Wiley and Sons, Inc. DOI 10.1002/9780470479216.corpsy0943, URL http://dx.doi.org/10.1002/9780470479216.corpsy0943
17. Tucker P (2003) The impact of rest breaks upon accident risk, fatigue and performance: a review. Work & Stress 17(2):123–137
18. Williamson RJ, Purcell S, Sterne A, Wessely S, Hotopf M, Farmer A, Sham PC (2005) The relationship of fatigue to mental and physical health in a community sample. Social psychiatry and psychiatric epidemiology 40(2):126–132